

~~0590
1124~~

Page 1 of 7



OIPE

RAW SEQUENCE LISTING DATE: 02/10/2002
PATENT APPLICATION: US/09/837,602 TIME: 13:38:33

Input Set : N:\CrF3\RULE60\09837602.txt
Output Set: N:\CRF3\02102002\I837602.raw

4 <110> APPLICANT: Petrini, John H.J.
5 Morgan, William Franklin
6 Maser, Richard Scott
7 Carney, James Patrick
10 <120> TITLE OF INVENTION: DNA Encoding A DNA Repair Protein
13 <130> FILE REFERENCE: 800.019US1
15 <140> CURRENT APPLICATION NUMBER: US 09/837,602
17 <141> CURRENT FILING DATE: 2001-04-18
19 <150> PRIOR APPLICATION NUMBER: US 09/067,641
21 <151> PRIOR FILING DATE: 1998-04-27
23 <160> NUMBER OF SEQ ID NOS: 24
25 <170> SOFTWARE: FastSEQ for Windows Version 3.0
27 <210> SEQ ID NO: 1
28 <211> LENGTH: 4403
29 <212> TYPE: DNA
30 <213> ORGANISM: Homo sapiens
32 <400> SEQUENCE: 1
33 ttcggcacga ggcgcgggttgcacgtcgcccccagccctgaggagccggac cga
34 actgctgccc gccgcggggcc cggcaggagg agaaccatac agacttttgatctg
35 gtacgttggtaaaaaactgtgcattctaattgaaatgtatcgtcgat
36 aaatcatgctgtgttaactgctaactttctgttaaccaacctgagtcaaa caga
37 ccctgttattgacattaaaatgtatctaaatgtatggtacc ttgttaataggat
38 gcagaatggctttccgaa ctttgaagtcggggatggattacttttgagat
39 aagtaaattcagaatagagtatgacgccttggtgcatgtcttcttgtttag
40 tggggaaactgctttaatc aagctatattgcaacttggaggattactgtaa
41 gacagaagaa tgcactcacc ttgtcatggatcagtggaaa gttaccatta aaaa
42 tgcaactcatttgtggacgtccaattgtaaa gccagaatatttactgaattcc
43 agttcagtc aagaagcagccctccacaaat tgaaagtttttacccaccccttgg
44 atctatttggaaatgaaaaatgttgcattgtcaggacggcag gaaagaaaaacaaa
45 aggaaaaaca ttttatatttttgaatgcacaaacagcataagaaattgagttccg
46 ctttggaggtggggaaagctatgttgataaacagaagagaatgaagaagaaacata
47 tttggctccggaaacgtgttgttgatcaggataacaactcacaagaccc
48 tgactgtcagaagaaatggatcagtcaataatggatatgttccaaaggc aagg
49 acctatttcttgaagcagaaatggattggcgttgattttcactactacaaagaa
50 tgatcctcaggccatccccatcaggattaaagacaacaactccaggaccaac
51 acaaggcgtgtcagttgtatgaaaaactaatgtccaaagcgcctcaggtaaca
52 cgtagctgacacagaatcagagcaagcaga tacatggatttgatgtggaaa
53 aatcaaagtc tccaaaatggaaacaaaaattcagaatgttgcattcacaagacgc
54 aaaggagtccatcggaaacaaatgtctataataatgtatgtatgtatcaat
55 gatgagaatccaaactatcagcttccacc aactaaatttgcaggatataa
56 agataggcttctcagcagcagcagaccaatccatcaga aactactttc
57 caaaaaaaaaagg gaaaggatgaaatcaagaaatgtctcatgaaat
58 aqaaaacqtcttggatggatggatggatggatggatggatggatggatggat

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/837,602

DATE: 02/10/2002
TIME: 13:38:33

Input Set : N:\Crf3\RULE60\09837602.txt
Output Set: N:\CRF3\02102002\I837602.raw

59	taaggagcag	catctatctg	agaatgagcc	tgtggacaca	aactcagaca	ataacttatt	1620
60	tacagataca	gattaaaaat	ctattgtgaa	aaattctgcc	agtaaatctc	atgctgcaga	1680
61	aaagctaaga	tcaaataaaa	aaagggaaat	ggatgatgtg	gccatagaag	atgaagtatt	1740
62	ggaacagtta	ttcaaggaca	aaaaaccaga	gttagaaatt	gatgtgaaag	ttcaaaaaca	1800
63	ggaggaagat	gtcaatgtt	aaaaaggcc	aaggatggat	atagaaacaa	atgacacttt	1860
64	cagtgtgaa	gcagtaccag	aaagtagcaa	aatatctcaa	aaaaatgaaa	ttggaaagaa	1920
65	acgtgaactc	aaggaagact	cactatggc	agctaaagaa	atatctaaca	atgacaaaact	1980
66	tcaggatgat	agtggatgc	ttccaaaaaa	gctgttattt	actgaattt	gatcactgg	2040
67	gattaaaaac	tctacttcca	gaaatccgtc	tggcataaat	gatgattatg	gtcaactaaa	2100
68	aaatttcaag	aaattcaaaa	aggtcacata	tcctggagca	ggaaaacttc	cacacatcat	2160
69	tggaggatca	gatctaata	ctcatcatgc	tcgaaaagaat	acagaactag	aagagtggct	2220
70	aaggcaggaa	atggaggtac	aaaatcaaca	tgcaaaagaa	gagtctttt	ctgtatgtat	2280
71	tttttagatac	aatccattt	taaaaaaggag	aagataactg	aggattttaa	aaagaagcca	2340
72	tggaaaaact	tccttagtaag	catctacttc	aggccaacaa	ggttatatga	atatatagtg	2400
73	tatagaagcg	atthaagttt	caatgtttt	tggcctaaat	ttattaaata	aaatgcacaa	2460
74	aactttgatt	cttttgtat	taacaattt	ttgttcttt	ttcaggctt	gtcattgcat	2520
75	cttttttca	tttttaatg	tgtttttt	attaaatagt	taatatagtc	acagttcaaa	2580
76	attctaaatr	tacgttaaggt	aaaggactaa	agtcaccctt	ccaccatgt	cctagctact	2640
77	tatttttaaa	taatttccca	cacaaatgt	agcataacat	atgcagtgtt	ctcacaccc	2700
78	cttttttact	tagtaagatt	aaaaattata	ggaatataaa	tataatgtt	ttaatatttt	2760
79	ttcttttcca	ttatgtgtt	gtcttaccta	aactctgg	atccaaacaa	aatggcttca	2820
80	gtgggtcaga	tgtcacctac	atgttattt	agtactagaa	actgaagacc	atgtggagac	2880
81	ttcatcaaac	atgggtttag	tttacccag	aatggaaaga	cctgtacccc	tttttgttgg	2940
82	tcttactgag	ctgggtgggt	gtctgtttt	agcttattt	gagtcctagt	tttcttactt	3000
83	ataaaatgt	aatggtgaga	ttgtttt	tttctacc	aaaggagat	ggtaagaaac	3060
84	aatgaatgtc	tttttcaaa	cttatttgc	aagtgtattt	caagtctgt	ttcaaaaata	3120
85	tattcatgt	cctgtatcc	agcaagaagg	gagttccagt	caagagtac	tacaactgt	3180
86	tagttttt	gagaatgaga	aatggacag	tgaggaatgg	aggccatatt	tccatgactt	3240
87	cccttgtaaa	cagaagcaac	agaagggaca	agaggctgg	ctctacatca	ctctcacctt	3300
88	ccaaatctt	tggaaatgt	tctacttgcc	agaaccaa	taacttactt	ccaaatctg	3360
89	gctgcttgc	ggtggactc	cagctgcaag	ggagtttaggg	aatgaaggt	ctttttttaa	3420
90	aagcttctca	gccttccat	ggaacagaaa	ttgggtgagc	caatctgca	tttctactac	3480
91	aggcatttgc	accagtttgc	ttattgaaat	attatagaga	gttatgaaca	cttaaattat	3540
92	gatagtggta	tgacatttgc	tagaacatgg	gatacttt	aagttagatt	gacagggcat	3600
93	attagttgt	gaaatggagt	catttgcgt	tyttaatagc	catgtatcat	aattaccaag	3660
94	tgaagctgg	ggaacatatg	gtctccattt	tacgtttaag	gaatataatg	gacagattaa	3720
95	tattgttgc	tgtcatgccc	acaatccctt	tctaaggaag	actgccctac	tatagcattt	3780
96	tttatatttt	tcaatttgc	aatataatg	atgaggagtt	ctggtaccc	ctgtctttac	3840
97	aaatatttgg	tgttgcct	tatttttcc	tttttaaccm	ttcccaattt	gggtgtgt	3900
98	gtggatgtt	ccatttgggt	ttaattttt	atatccctg	tagctataat	tgggtcatag	3960
99	aaattcttta	tacattctag	atgcaagtt	cttgcggat	atacgtattt	agatattaca	4020
100	cctagtctgt	ggcttgcgt	ttttctt	gtcttttgc	aatagaagt	tttaaatttt	4080
101	gacaaggctca	aatttatttt	tttctttt	ttgatatttt	ttctctccaa	tttaacccca	4140
102	agatttcaga	tatttgcgt	tattatataa	actttatattt	tttatattt	tgtatctac	4200
103	tgaattgata	tgtatgtt	gaattatgg	tcagggttct	ttttttcccc	catacaagta	4260
104	tccagtcatt	gtaacactgt	ttattgaaag	aattatcc	tcctcattaa	attaccc	4320
105	caatttagtaa	aaaatcaatt	aaccatrmr	mmrrrggat	ccactagttc	tagagcggcc	4380
106	gccaccgcgg	tggagctcca	gct				4403
108	<210>	SEQ ID NO:	2				

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/837,602

DATE: 02/10/2002
TIME: 13:38:33

Input Set : N:\Crf3\RULE60\09837602.txt
Output Set: N:\CRF3\02102002\I837602.raw

109 <211> LENGTH: 754
110 <212> TYPE: PRT
111 <213> ORGANISM: Homo sapiens
113 <400> SEQUENCE: 2
114 Met Trp Lys Leu Leu Pro Ala Ala Gly Pro Ala Gly Gly Glu Pro Tyr
115 1 5 10 15
116 Arg Leu Leu Thr Gly Val Glu Tyr Val Val Gly Arg Lys Asn Cys Ala
117 20 25 30
118 Ile Leu Ile Glu Asn Asp Gln Ser Ile Ser Arg Asn His Ala Val Leu
119 35 40 45
120 Thr Ala Asn Phe Ser Val Thr Asn Leu Ser Gln Thr Asp Glu Ile Pro
121 50 55 60
122 Val Leu Thr Leu Lys Asp Asn Ser Lys Tyr Gly Thr Phe Val Asn Glu
123 65 70 75 80
124 Glu Lys Met Gln Asn Gly Phe Ser Arg Thr Leu Lys Ser Gly Asp Gly
125 85 90 95
126 Ile Thr Phe Gly Val Phe Gly Ser Lys Phe Arg Ile Glu Tyr Glu Pro
127 100 105 110
128 Leu Val Ala Cys Ser Ser Cys Leu Asp Val Ser Gly Lys Thr Ala Leu
129 115 120 125
130 Asn Gln Ala Ile Leu Gln Leu Gly Gly Phe Thr Val Asn Asn Trp Thr
131 130 135 140
132 Glu Glu Cys Thr His Leu Val Met Val Ser Val Lys Val Thr Ile Lys
133 145 150 155 160
134 Thr Ile Cys Ala Leu Ile Cys Gly Arg Pro Ile Val Lys Pro Glu Tyr
135 165 170 175
136 Phe Thr Glu Phe Leu Lys Ala Val Gln Ser Lys Lys Gln Pro Pro Gln
137 180 185 190
138 Ile Glu Ser Phe Tyr Pro Pro Leu Asp Glu Pro Ser Ile Gly Ser Lys
139 195 200 205
140 Asn Val Asp Leu Ser Gly Arg Gln Glu Arg Lys Gln Ile Phe Lys Gly
141 210 215 220
142 Lys Thr Phe Ile Phe Leu Asn Ala Lys Gln His Lys Lys Leu Ser Ser
143 225 230 235 240
144 Ala Val Val Phe Gly Gly Glu Ala Arg Leu Ile Thr Glu Glu Asn
145 245 250 255
146 Glu Glu Glu His Asn Phe Phe Leu Ala Pro Gly Thr Cys Val Val Asp
147 260 265 270
148 Thr Gly Ile Thr Asn Ser Gln Thr Leu Ile Pro Asp Cys Gln Lys Lys
149 275 280 285
150 Trp Ile Gln Ser Ile Met Asp Met Leu Gln Arg Gln Gly Leu Arg Pro
151 290 295 300
152 Ile Pro Glu Ala Glu Ile Gly Leu Ala Val Ile Phe Met Thr Thr Lys
153 305 310 315 320
154 Asn Tyr Cys Asp Pro Gln Gly His Pro Ser Thr Gly Leu Lys Thr Thr
155 325 330 335
156 Thr Pro Gly Pro Ser Leu Ser Gln Gly Val Ser Val Asp Glu Lys Leu
157 340 345 350
158 Met Pro Ser Ala Pro Val Ash Thr Thr Tyr Val Ala Asp Thr Glu

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/837,602

DATE: 02/10/2002
TIME: 13:38:33

Input Set : N:\Crf3\RULE60\09837602.txt
Output Set: N:\CRF3\02102002\I837602.raw

159	355	360	365
160	Ser Glu Gln Ala Asp Thr Trp Asp Leu Ser Glu Arg Pro Lys Glu Ile		
161	370	375	380
162	Lys Val Ser Lys Met Glu Gln Lys Phe Arg Met Leu Ser Gln Asp Ala		
163	385	390	395
164	400		
165	Pro Thr Val Lys Glu Ser Cys Lys Thr Ser Ser Asn Asn Asn Ser Met		
166	405	410	415
167	Val Ser Asn Thr Leu Ala Lys Met Arg Ile Pro Asn Tyr Gln Leu Ser		
168	420	425	430
169	Pro Thr Lys Leu Pro Ser Ile Asn Lys Ser Lys Asp Arg Ala Ser Gln		
170	435	440	445
171	Gln Gln Gln Thr Asn Ser Ile Arg Asn Tyr Phe Gln Pro Ser Thr Lys		
172	450	455	460
173	Lys Arg Glu Arg Asp Glu Glu Asn Gln Glu Met Ser Ser Cys Lys Ser		
174	465	470	475
175	480		
176	Ala Arg Ile Glu Thr Ser Cys Ser Leu Leu Glu Gln Thr Gln Pro Ala		
177	485	490	495
178	Thr Pro Ser Leu Trp Lys Asn Lys Glu Gln His Leu Ser Glu Asn Glu		
179	500	505	510
180	Pro Val Asp Thr Asn Ser Asp Asn Asn Leu Phe Thr Asp Thr Asp Leu		
181	515	520	525
182	Lys Ser Ile Val Lys Asn Ser Ala Ser Lys Ser His Ala Ala Glu Lys		
183	530	535	540
184	Leu Arg Ser Asn Lys Lys Arg Glu Met Asp Asp Val Ala Ile Glu Asp		
185	545	550	555
186	560		
187	Glu Val Leu Glu Gln Leu Phe Lys Asp Thr Lys Pro Glu Leu Glu Ile		
188	565	570	575
189	Asp Val Lys Val Gln Lys Gln Glu Glu Asp Val Asn Val Arg Lys Arg		
190	580	585	590
191	Pro Arg Met Asp Ile Glu Thr Asn Asp Thr Phe Ser Asp Glu Ala Val		
192	595	600	605
193	Pro Glu Ser Ser Lys Ile Ser Gln Glu Asn Glu Ile Gly Lys Lys Arg		
194	610	615	620
195	Glu Leu Lys Glu Asp Ser Leu Trp Ser Ala Lys Glu Ile Ser Asn Asn		
196	625	630	635
197	640		
198	Asp Lys Leu Gln Asp Asp Ser Glu Met Leu Pro Lys Lys Leu Leu		
199	645	650	655
200	Thr Glu Phe Arg Ser Leu Val Ile Lys Asn Ser Thr Ser Arg Asn Pro		
201	660	665	670
202	Ser Gly Ile Asn Asp Asp Tyr Gly Gln Leu Lys Asn Phe Lys Lys Phe		
203	675	680	685
204	Lys Lys Val Thr Tyr Pro Gly Ala Gly Lys Leu Pro His Ile Ile Gly		
205	690	695	700
206	Gly Ser Asp Leu Ile Ala His His Ala Arg Lys Asn Thr Glu Leu Glu		
207	705	710	715
208	Glu Trp Leu Arg Gln Glu Met Glu Val Gln Asn Gln His Ala Lys Glu		
209	725	730	735
210	Glu Ser Leu Ala Asp Asp Leu Phe Arg Tyr Asn Pro Tyr Leu Lys Arg		
211	740	745	750

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/837,602

DATE: 02/10/2002
TIME: 13:38:33

Input Set : N:\Crf3\RULE60\09837602.txt
Output Set: N:\CRF3\02102002\I837602.raw

208 Arg Arg
211 <210> SEQ ID NO: 3
212 <211> LENGTH: 87
213 <212> TYPE: PRT
214 <213> ORGANISM: Homo sapiens
216 <220> FEATURE:
217 <221> NAME/KEY: UNSURE
218 <222> LOCATION: (48)...(48)
219 <223> OTHER INFORMATION: Unsure
221 <400> SEQUENCE: 3
222 Tyr Val Val Gly Arg Lys Asn Cys Ala Ile Leu Ile Glu Asn Asp Gln
223 1 5 10 15
224 Ser Ile Ser Arg Asn His Ala Val Leu Thr Ala Asn Phe Ser Val Thr
225 20 25 30
W--> 226 Asn Leu Ser Gln Thr Asp Glu Ile Pro Val Leu Thr Leu Lys Asn Xaa
227 35 40 45
228 Lys Tyr Gly Thr Phe Val Asn Glu Glu Lys Met Gln Asn Gly Phe Ser
229 50 55 60
230 Arg Thr Leu Lys Ser Val Asp Gly Ile Thr Phe Gly Val Phe Gly Ser
231 65 70 75 80
232 Lys Phe Arg Ile Glu Tyr Glu
233 85
235 <210> SEQ ID NO: 4
236 <211> LENGTH: 87
237 <212> TYPE: PRT
238 <213> ORGANISM: Homo sapiens
240 <400> SEQUENCE: 4
241 Tyr Ser Ile Gly Arg Ser Ser Lys Asn Pro Leu Ile Ile Lys Asn Asp
242 1 5 10 15
243 Lys Ser Ile Ser Arg Gln His Ile Thr Phe Lys Trp Glu Ile Asn Asn
244 20 25 30
245 Ser Ser Asp Leu Lys His Ser Ser Leu Cys Leu Val Asn Lys Gly Lys
246 35 40 45
247 Leu Thr Ser Leu Asn Lys Lys Phe Met Lys Val Gly Glu Thr Phe Thr
248 50 55 60
249 Ile Asn Ala Ser Cys Val Leu Lys Ser Thr Ile Glu Leu Gly Thr Thr
250 65 70 75 80
251 Pro Ile Arg Ile Glu Phe Glu
252 85
254 <210> SEQ ID NO: 5
255 <211> LENGTH: 13
256 <212> TYPE: PRT
257 <213> ORGANISM: Homo sapiens
259 <400> SEQUENCE: 5
260 Asn Pro Ser Gly Leu Asn Asp Asp Tyr Gly Gln Leu Lys
261 1 5 10
263 <210> SEQ ID NO: 6
264 <211> LENGTH: 680
265 <212> TYPE: PRT

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/837,602

DATE: 02/10/2002
TIME: 13:38:34

Input Set : N:\Crf3\RULE60\09837602.txt
Output Set: N:\CRF3\02102002\I837602.raw

L:226 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3